Structure, Vol. 9, 125–132, February, 2001, ©2001 Elsevier Science Ltd. All rights reserved. PII S0969-2126(01)00566-4

Crystal Structure of the 100 kDa Arsenite Oxidase from *Alcaligenes faecalis* in Two Crystal Forms at 1.64 Å and 2.03 Å

Paul J. Ellis,* Thomas Conrads,† Russ Hille,† and Peter Kuhn*‡

*Stanford Synchrotron Radiation Laboratory
Stanford University
Stanford, California 94309

†Department of Molecular and Cellular Biochemistry
The Ohio State University
Columbus, Ohio 43210

Summary

Background: Arsenite oxidase from *Alcaligenes faecalis* NCIB 8687 is a molybdenum/iron protein involved in the detoxification of arsenic. It is induced by the presence of AsO_2^- (arsenite) and functions to oxidize $As^{III}O_2^-$, which binds to essential sulfhydryl groups of proteins and dithiols, to the relatively less toxic $As^{V}O_4^{3-}$ (arsenate) prior to methylation.

Results: Using a combination of multiple isomorphous replacement with anomalous scattering (MIRAS) and multiple-wavelength anomalous dispersion (MAD) methods, the crystal structure of arsenite oxidase was determined to 2.03 Å in a $P2_1$ crystal form with two molecules in the asymmetric unit and to 1.64 Å in a P1 crystal form with four molecules in the asymmetric unit. Arsenite oxidase consists of a large subunit of 825 residues and a small subunit of approximately 134 residues. The large subunit contains a Mo site, consisting of a Mo atom bound to two pterin cofactors, and a [3Fe-4S] cluster. The small subunit contains a Rieske-type [2Fe-2S] site.

Conclusions: The large subunit of arsenite oxidase is similar to other members of the dimethylsulfoxide (DMSO) reductase family of molybdenum enzymes, particularly the dissimilatory periplasmic nitrate reductase from *Desulfovibrio desulfuricans*, but is unique in having no covalent bond between the polypeptide and the Mo atom. The small subunit has no counterpart among known Mo protein structures but is homologous to the Rieske [2Fe-2S] protein domain of the cytochrome bc_1 and cytochrome b_6f complexes and to the Rieske domain of naphthalene 1,2-dioxygenase.

Introduction

Arsenic is broadly distributed in the environment, arising from both natural and anthropogenic processes. Most specific chemical forms of arsenic, including both arsenite ($As^{III}O_2^-$) and arsenate ($As^{V}O_4^{3-}$), are toxic to some degree. Arsenite binds to sulfhydryl groups of proteins and dithiols such as glutaredoxin, disrupting intracellular oxidation–reduction homeostasis. The toxicity of arsenate is derived from it being a structural analog of phosphate; ATPase will arsenylate ADP, which subsequently spontaneously hydrolyzes to create a futile cycle within the cell [1]. Bacteria overcome the toxic effects of arsenite and arsenate by either actively exporting them or by chemically modifying them to less toxic forms. Arsenite can either be methylated directly or oxidized to the relatively less toxic arsenate prior to methylation [1].

Arsenite oxidase from the soil pseudomonad *Alcaligenes faecalis* is a 100 kDa molybdenum- and iron-sulfur-containing protein located on the outer surface of the inner membrane of this Gram-negative organism [1]. It is arsenite inducible and functions to oxidize arsenite to arsenate according to the following reaction:

$$As^{III}O_2^- + 2H_2O \rightarrow As^{V}O_4^{3-} + 4H^+ + 2e^-$$

Reducing equivalents generated in the course of the reaction are subsequently transferred to the periplasmic electron acceptors azurin and cytochrome c to complete the reaction sequence [1].

Although arsenite oxidase is thought to possess two equivalents of the pyranopterin cofactor common to all molybdenum-containing enzymes other than nitrogenase and can thus be assigned to the dimethylsulfoxide (DMSO) reductase family of enzymes [2, 3], it has some unusual properties compared to the other members of this family for which crystal structures are known. In particular, it has been shown by electron paramagnetic resonance (EPR) spectroscopy to possess either a HiPIP [4Fe-4S] or [3Fe-4S] center and a Rieske-type [2Fe-2S] center. These centers are not unreasonable in that the catalyzed reaction occurs with a potential of approximately +60 mV [4] and these iron centers are known to typically have reduction potentials above 0 mV. Nevertheless, the constitution of arsenite oxidase contrasts with that of the five other members for which there are structures, in which two possess only a single [4Fe-4S] site (formate dehydrogenase H of *E. coli* and the Nap nitrate reductase of *Desulfovibrio desulfuricans*), while the remaining members (DMSO reductase from *Rhodobacter sphaeroides* or *capsulatus* and trimethylamine-*N*-oxide reductase from *Shewanella massilia*) have no metal cofactors apart from the molybdenum center [1, 5, 6, 7, 8, 9].

We have determined the structure of *A. faecalis* arsenite oxidase in two crystal forms. Here we give a description of arsenite oxidase and compare it to other related molecules.

Results and Discussion

Structure Determination

The structure of arsenite oxidase was determined using a combination of multiple-isomorphous replacement

‡To whom correspondence should be addressed (e-mail: pkuhn@stanford.edu).

Key words: arsenite oxidase; [3Fe-4S] cluster; [2Fe-2S] cluster; crystal structure; molybdopterin cofactor; multiwavelength anomalous dispersion

with anomalous scattering (MIRAS) and multiple-wavelength anomalous dispersion (MAD) techniques using data from a P1 form with four molecules in the asymmetric unit and a $P2_1$ form with two molecules in the asymmetric unit (see Experimental Procedures).

The P1 and $P2_1$ forms were refined to 1.64 Å and 2.03 Å, respectively. As the amino acid sequence was not initially known, each residue was assigned manually during the process of refinement on the basis of map density and hydrogen bonding (see Experimental Procedures). Subsequently, the model was corrected using a partial sequence comprising approximately 70% of the protein chain. Although the remaining 30%, or 284 residues, is expected to incorporate approximately 30 incorrect assignments, no residues referred to in the following discussion are implicated.

Although at somewhat different nominal pH, 6.4 versus. 8.5, respectively, and despite marked differences in crystal packing, the P1 and $P2_1$ structures are very similar. Neglecting several residues at the N and C termini of the chains, the rmsd of the backbone atoms (N, CA, and C) of the average $P2_1$ structure from the corresponding atoms of the average P1 structure is just 0.7 Å. Unless otherwise noted, the following discussion refers to the higher resolution P1 structure only.

Structure Description

A. faecalis arsenite oxidase is a heterodimer consisting of a large subunit of 825 residues, incorporating the molybdenum center and a [3Fe-4S] cluster, and a small, or "Rieske", subunit of approximately134 residues, incorporating a Rieske-type [2Fe-2S] cluster.

The two subunits are intimately associated, with the interface burying approximately 3790 $Å^2$ of accessible surface area and including four backbone–backbone hydrogen bonds and approximately 27 other direct hydrogen bonds. Together, they form an irregularly shaped complex with overall dimensions of approximately 75 Å × 75 Å × 50 Å (Figure 1).

Large Subunit

The large subunit of arsenite oxidase is divided into four domains (Figure 1), with domain I binding the [3Fe-4S] cluster and the Rieske subunit and domains II–IV binding the molybdenum center. Domain I, comprising residues 1–119, 532–558, and 622–657, consists of three antiparallel β sheets and six helices. The [3Fe-4S] cluster is coordinated by the motif Cys21-X_2-Cys24-X_3-Cys28 near the interface with domains III and IV. Domain II, residues 120–195, 424–531, and 559–621, and domain III, residues 196–423 and 658–682, have similar αβα sandwich topologies, with a large β sheet packed by helices on either side. Domain II consists of a seven-stranded, mainly parallel β sheet with five helices on one side and seven helices and a small, antiparallel, two-stranded β sheet on the other. Domain III consists of a fully parallel, five-stranded β sheet with six helices on one side and nine helices and a small, two-stranded, parallel β sheet on the other. Domain IV, residues 683–825, is a six-stranded β barrel capped on both ends with a helix and flanked on one side by a further helix and on the other side by three more.

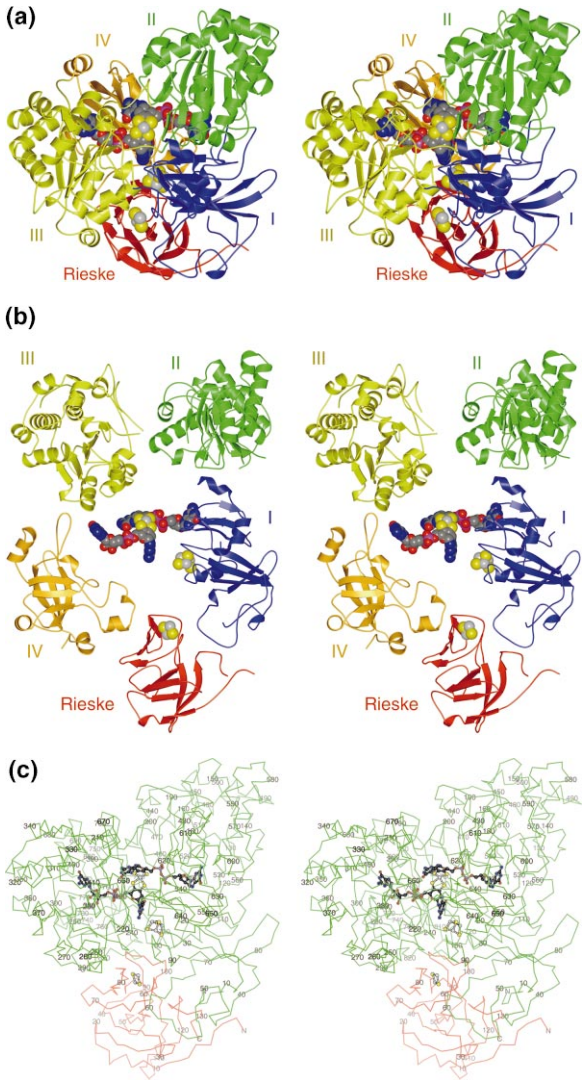

Figure 1. Stereo Ribbon Diagram of Arsenite Oxidase Looking Down onto the Active Site

(a) Domains I–IV of the large subunit are drawn in blue, green, yellow, and orange, respectively. The Rieske subunit is drawn in red. The molybdenum cofactor, [3Fe-4S] cluster, and [2Fe-2S] cluster are also shown. The oxo group coordinated to the molybdenum atom has been omitted for clarity.

(b) An exploded view of arsenite oxidase showing the individual domains.

(c) A Cα trace of arsenite oxidase, with the large subunit in green and the Rieske subunit in red. Every tenth amino acid is labeled. The figures were prepared using the programs MOLSCRIPT [31] and POV-Ray.

Molybdenum Center

A large, flattened funnel-like cavity bounded by domains I, II, and III leads to the molybdenum center located near the center of the molecule (Figure 1). The Mo atom is symmetrically coordinated to the four *cis*-dithiolene sulfur atoms of two antiparallel molybdopterin guanosine dinuleotide (MGD) cofactors that are firmly bound by a network of hydrogen bonds and salt bridges in the space between the domains of the large subunit (Figure 2). One MGD cofactor, conventionally labeled P, lies mainly between domains II and IV, and the other, labeled Q,

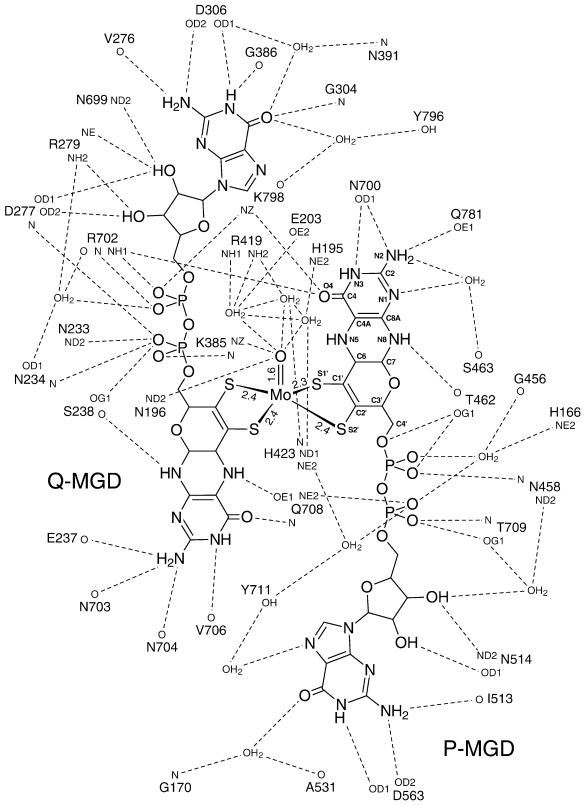

Figure 2. Schematic Representation of the Hydrogen-Bonding Interactions of the Molybdenum Cofactor

The average Mo-S and Mo=O distances in the P1 form are indicated. Hydrogen bonds were identified using the program HBPLUS [32].

lies between domains III and IV. Both domains II and III resemble the classical nucleotide binding fold [10] and exhibit a marked pseudosymmetry relative to each other, reflecting the pseudosymmetry of the molybdenum cofactor.

In both the P1 and $P2_1$ structures, electron-density maps show a fifth ligand, modeled as a single O atom (Figure 3a). The short refined Mo–O distance of 1.6 Å in both the P1 and $P2_1$ structures suggests a Mo=O bond. The overall coordination geometry is approximately square pyramidal, with the Mo atom displaced approximately 0.8 Å from the mean dithiolene sulfur plane toward the axial oxygen atom.

On the basis of a comparison with other members of this family of molybdenum enzymes, and consistent with recent X-ray absorption studies of arsenite oxidase (George et al., unpublished data), the five-coordinate geometry of the molybdenum suggests that the protein has become reduced in the course of crystallographic data acquisition. Photoreduction of the metal centers was also suggested by a shift of the Fe absorption edge toward lower energy over the course of the MAD data collection. In subsequent tests using P1 crystals, the energy of the Fe edge was also observed to decrease by approximately 3 eV over a period of 20 min of X-ray exposure. When a crystal was removed from the X-ray beam, the energy of the edge slowly increased with time, suggesting that the Fe centers were becoming reoxidized and that during data collection, the crystal

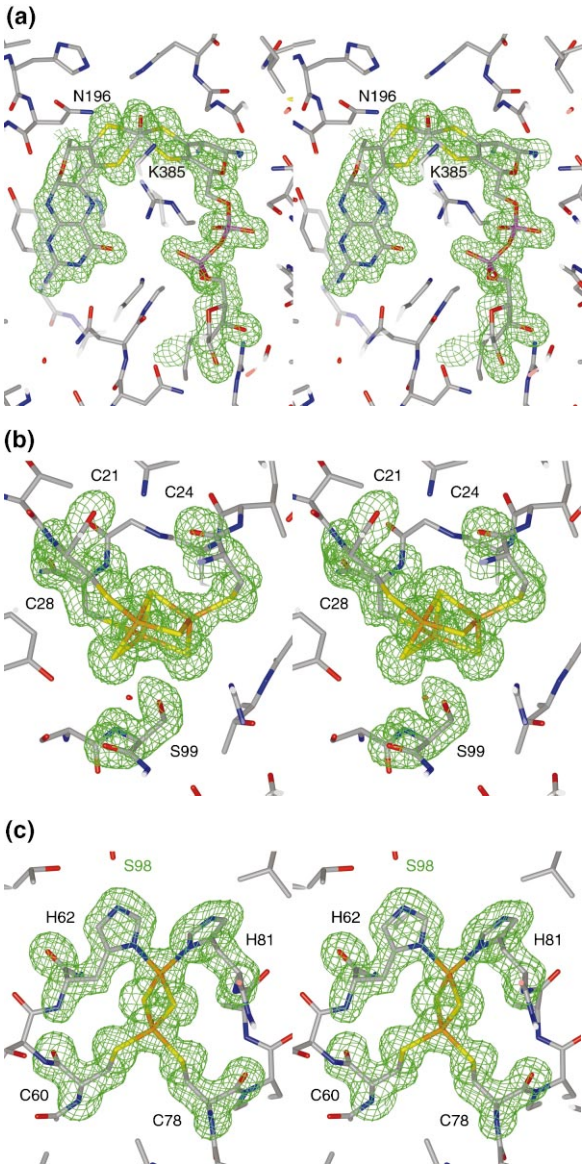

Figure 3. Stereo Diagram of Annealed σ_A-Weighted F_o-F_c Omit Maps for One of the Four Molecules in the P1 Cell

(a) Molybdenum center.
(b) [3Fe-4S] center.
(c) [2Fe-2S] center of one of the 4 molecules in the P1 cell.
All maps are contoured at the 4σ level. The figures were prepared using the programs CNS [26], XFIT [27], and POV-Ray.

is in fact an equilibrium mixture of oxidation states. However, if the molybdenum center is six-coordinate in the oxidized state and five-coordinate in the reduced state, then at least as far as this site is concerned, the equilibrium must lie strongly toward the reduced side under the conditions of data collection, as the metal appears clearly five-coordinate in both of the final P1 and $P2_1$ electron-density maps.

[3Fe-4S] Cluster

The [3Fe-4S] cluster is coordinated by residues of domain I, approximately 14 Å from the Mo atom and 15 Å from the [2Fe-2S] cluster of the Rieske subunit. Struc-

ture-based sequence alignment indicates that the sequence Cys-X_2-Cys-X_3-Cys-X_{26}-Cys, binding the corresponding [4Fe-4S] cluster in the dissimilatory nitrate reductase [6] and formate dehydrogenase H [5], is Cys21-X_2-Cys24-X_3-Cys28-X_{70}-Ser99 in arsenite oxidase (Figure 3b). The serine residue retains the position and orientation of the fourth cysteine and appears to play a role in electron transfer from the [3Fe-4S] cluster to the Rieske [2Fe-2S] center (see below).

Rieske Subunit

The Rieske subunit of arsenite oxidase can be divided into two subdomains, with an incomplete six-stranded, antiparallel β barrel at one end of the protein (β2-β3-β4-β9-β10-β1), and the cluster binding subdomain consisting of the loop between β4 and β5 and a four-stranded antiparallel β sheet (β5-β6-β7-β8) at the other (Figure 1b). The Rieske-type [2Fe-2S] cluster is coordinated by Cys60 and His62 from the β4-β5 loop and by Cys78 and His81 from the β6-β7 loop (Figure 3c). A disulfide bridge between Cys65 and Cys80 holds the two loops together. The imidazole rings of His62 and His81 coordinate the Fe atom closest to the end of the protein and are exposed to the surface of the subunit. In the complex, His62 is buried by the large subunit so that only His81 is exposed to solvent. Cys60 and Cys78 coordinate the other Fe atom and are buried within the protein.

Comparison with Other Mo Enzyme Structures

Although the large subunit of arsenite oxidase has relatively little sequence homology compared to the five other members of the DMSO reductase (DMSOR) family of molybdenum enzymes for which structures have been determined, the overall topology is similar both with respect to the division into four domains and to the fold of the individual domains. The most significant difference between arsenite oxidase and the other proteins is the absence of any covalent linkage between the protein and the molybdenum atom. As mentioned earlier, the large subunit also differs from the two previous examples incorporating a [4Fe-4S] cluster, with a serine replacing the last cysteine residue in the cluster binding motif to produce a high-potential [3Fe-4S] site.

Of the members of the DMSO reductase family for which structures have been determined, the large subunit of arsenite oxidase is most similar to the dissimilatory nitrate reductase (NAP) from *Desulfovibrio desulfuricans* [6], with a sequence identity calculated using the program CE [11] of approximately 23%, and to the formate dehydrogenase (FDH) from *E. coli* [5], with an identity of 20%. Both of these proteins have similar cofactors to those in the large subunit of arsenite oxidase, with a single [4Fe-4S] cluster found in place of the [3Fe-4S] cluster in the present structure. As expected, the large subunit of arsenite oxidase is slightly less homologous to the three proteins with no Fe-S cofactors: DMSOR from *Rhodobacter capsulatus* [8], DMSOR from *R. sphaeroides* [7], and trimethylamine *N*-oxide reductase (TMNOR) from *Shewanella massilia* [9], with sequence identities of 18%, 16%, and 15%, respectively. The above mentioned division also holds for the structural similarity. As calculated using CE, the rmsd from

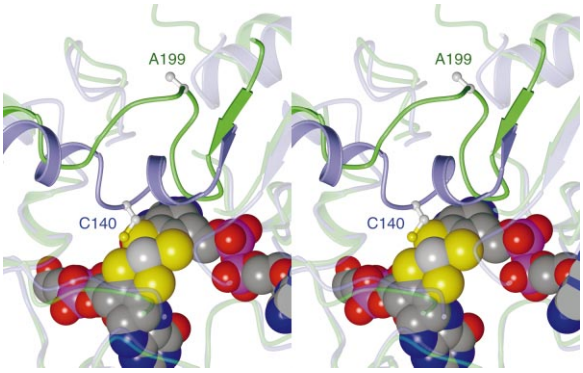

Figure 4. Stereo Ribbon Diagram of the Active Site of Arsenite Oxidase, Shown in Green, Superposed on the Active Site of NAP, Shown in Blue

For greater clarity, only the molybdenum center of arsenite oxidase is included and the molecules have been clipped. Also shown is one end of the Rieske domain of arsenite oxidase in red and the molybdenum center. The cysteine side chain coordinating the Mo atom in NAP lies near the metal atom. The corresponding alanine residue in arsenite oxidase is far from the active site. The figure was prepared using the programs MOLSCRIPT [31] and POV-Ray.

arsenite oxidase is 2.4 Å and 2.6 Å for NAP and FDH, respectively; 3.7 Å and 3.4 Å for the DMSO reductases from *R. capsulatus* and *R. sphaeroides*, respectively; and 3.3 Å for TMNOR. A more sensitive score is the structural diversity calculated using the program TOP3D [12]. NAP and FDH are most like the large subunit, with scores of 3.8 and 5.4, respectively, and *R. capsulatus* DMSOR, *R. sphaeroides* DMSOR, and TMNOR are considerably less homologous, with scores of 14, 12, and 12, respectively.

In all previously reported structures with a bis(molybdopterin) molybdenum site, the Mo atom is coordinated by the side chain of a serine, cysteine, or selenocysteine residue. In arsenite oxidase, the corresponding residue is Ala199 and there is no covalent bond between the Mo atom and the polypeptide chain (Figure 4). As a consequence of this difference, the loop containing this residue is folded away from the Mo site rather than toward it as in the previous structures (Figure 4), and the active site is significantly more exposed than in the prior examples.

Comparison with Other Rieske-Domain Structures

The Rieske [2Fe-2S] subunit of arsenite oxidase is unique among known Mo protein structures but is similar to the Rieske domains of cytochrome bc_1 [13, 14], cytochrome b_6f [15], and naphthalene 1,2-dioxygenase [16]. Figure 5 shows a superposition of the four known Rieske-domain structures. As observed previously [16], although all four Rieske domains retain the basic β barrel/β sandwich topology, the cluster binding subdomain is much more strongly conserved than the rest of the protein, with the cluster binding motif C-X-H-X_{15}-C-X_2-H in arsenite oxidase conforming to the known consensus sequence C-X-H-$X_{15–17}$-C-X_2-H.

Reaction Mechanism

On the basis of the present crystal structure, it is possible to draw several conclusions regarding the manner

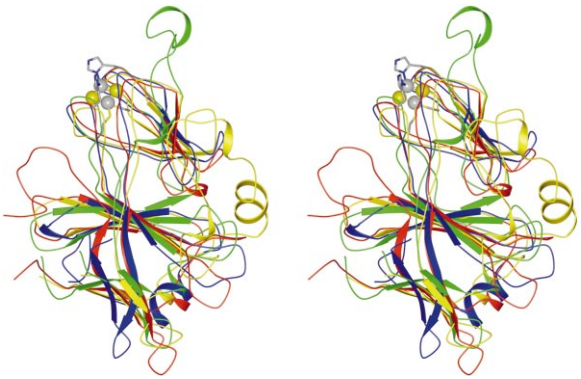

Figure 5. Stereo Ribbon Diagram of the Rieske Domain of Arsenite Oxidase, Shown in Red, Superposed on the Rieske Domains of Naphthalene 1,2-dioxygenase from *P. putida*, Cytochrome b_6f from Spinach, and Bovine Cytochrome bc_1, Shown in Green, Blue, and Yellow, Respectively

Also shown is the [2Fe-2S] cluster of arsenite oxidase and the cluster binding side chains. The figure was prepared using the programs MOLSCRIPT [31] and POV-Ray.

in which arsenite oxidase catalyzes the oxidation of arsenite. Solvent access to the molybdenum center is provided by a hydrophilic channel. Near the base of this channel is a cluster of amino acid residues, including His195, Glu203, Arg419, and His423. A manual fit of arsenite [as $AsO_2(OH)^{2-}$] suggests that these residues represent the substrate binding site of the enzyme. By analogy to other members of this family of enzymes and consistent with XAS and resonance Raman studies of oxidized arsenite oxidase (George et al., unpublished data), the oxidized form of the molybdenum center most likely possesses a terminal Mo=O group. Binding of arsenite by the indicated residues positions the As lone pair of electrons appropriately for nucleophilic attack on the Mo=O group of the oxidized molybdenum center (Figure 6). This reaction yields reduced molybdenum(IV) with product (arsenate) coordinated to the metal. The reductive half-reaction of the catalytic cycle is completed by dissociation of product from the molybdenum

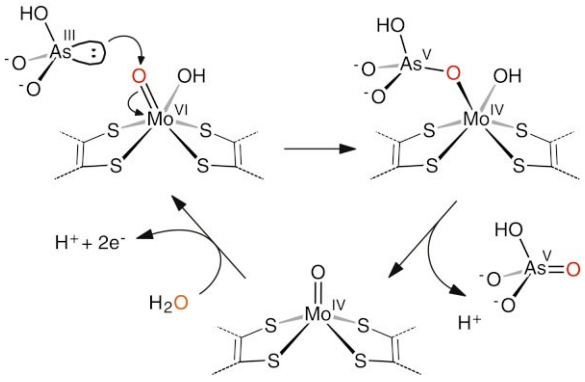

Figure 6. Schematic Representation of the Proposed Reaction Mechanism for the Oxidation of Arsenite by Arsenite Oxidase

The oxygen atom transferred from water to arsenite during the catalytic cycle is highlighted in red.

coordination sphere to give the reduced enzyme in the crystal structure.

The obligatory two-electron nature of the chemistry of this portion of the catalytic sequence provides a possible explanation as to why no Mo(V) EPR signal is observed in the course of the reaction of enzyme with arsenite [1]. The catalytic cycle is completed in the oxidative half-reaction, which involves reoxidation, with the addition and deprotonation of water to regenerate the oxo group, which returns to the starting oxidized enzyme. This chemistry is directly related to that exhibited by DMSO reductase and other members of this family of bacterial enzymes for which there is chemical precedent [2] and is specifically consistent with the observation that under single-turnover conditions in [18-O] water, 18-O is incorporated into the molybdenum center as a terminal oxo (Mo=O) group (George et al., unpublished data).

Electron-Transfer Pathway

Based on the arrangement of the metal clusters in arsenite oxidase, electrons generated by the oxidation of AsO_2^- follow the path Mo→[3Fe-4S]→[2Fe-2S]. From the Rieske [2Fe-2S] cluster, the electrons are then transferred to a periplasmic electron acceptor, either azurin or cytochrome c [1].

The nearest Fe atom of the [3Fe-4S] cluster lies approximately 12 Å from the Mo atom but less than 7 Å from the partly conjugated pterin pyrazine ring system of the Q-MGD. There are several possible electron-transfer pathways from the pyrazine system to the [3Fe-4S] cluster (Figure 7). The shortest of these is from the pterin N8 atom to the carbonyl O atom of Ser238 and then through space to SG of Cys24 (N8–O = 2.8 Å, O–SG = 3.8 Å). A longer pathway, mediated only by covalent and hydrogen bonds, is from the NH_2 group via the carbonyl O of Asn704 to the side chain of Arg101 and then to the S1 atom of the cluster (NH_2–O = 3.1 Å, O–NH1 Arg101 = 2.8 Å, and Arg101 NE–S1 = 3.4 Å).

Electron transfer from the [3Fe-4S] cluster to the [2Fe-2S] cluster is likewise partly mediated by hydrogen bonds. In this case, the shortest path is from the S3 atom of the [3Fe-4S] cluster to the amide N atom of Ser99, through the conjugated amide bond to the carbonyl O atom of Ser98, from the O atom across the subunit interface to the imidazole NE2 atom of the Rieske cluster binding residue His62, and finally to the [2Fe-2S] FE1 atom (S3–N = 3.4 Å, O–NE2 = 2.9 Å). Other potential pathways are from S1 or S3 of the [3Fe-4S] cluster to OG of Ser99 (S1–OG = 3.4 Å, S3–OG = 3.4 Å), through bonds to the amide group, and then to the [2Fe-2S] site as before.

The linear electron-transfer pathway implied by the present crystal structure is consistent with the observed structures of the two iron-sulfur centers, as these have reduction potentials of +300 mV or higher. The relative magnitudes of the reduction potentials of the two iron-sulfur centers are not known at present, but we emphasize that it need not be the case that the Rieske center, clearly the terminus of electron transfer on the basis of the crystal structure, have the higher potential. In the nickel-iron hydrogenase of *Desulfovibrio gigas*, for ex-

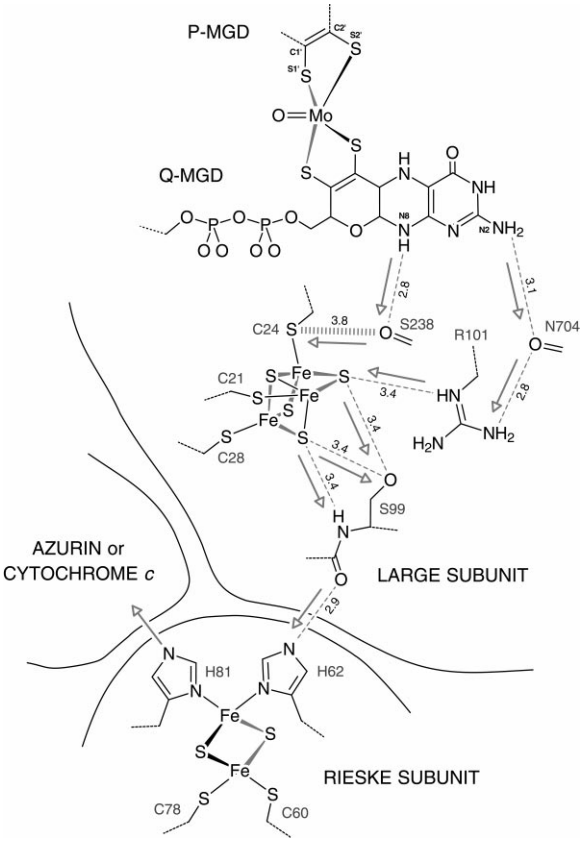

Figure 7. Schematic Representation of the Proposed Electron-Transfer Pathways from the Molybdenum to Azurin or Cytochrome c

ample, a linear sequence of four iron–sulfur centers is found in the crystal structure, with one of the internal ones being a [3Fe-4S] cluster of unusually high potential [17]. It has been noted, in fact, that so long as the distance over which electron transfer occurs is shorter than 12–15 Å, even "uphill" potentials as great as 300 mV are not likely to constitute barriers to catalysis [18].

From the [2Fe-2S] cluster, electrons are transferred either to azurin or cytochrome c [1], probably through the imidazole ring of His81. This role for His81 is suggested by examination of the arsenite oxidase structure in isolation: of the 4 cluster binding side chains, only that of His81 is solvent exposed, forming a hydrogen bond to a solvent molecule (NE2–O = 2.8 Å), and by analogy with the Rieske domain of the bovine-mitochondrial cytochrome bc_1 complex. In the "c_1" state, the Rieske subunit is oriented to form a hydrogen bond from the imidazole ring of the corresponding residue, His161, to a propionate group of the cytochrome c_1 heme [13]. When the Rieske center of cytochrome bc_1 in the "c_1" state is superposed on the arsenite oxidase Rieske center, the mitochondrial cytochrome lies mainly in the space between domains I and III of the large subunit. Although there is strong hindrance between arsenite oxidase and the cytochrome so that this particular protein, like horse heart cytochrome c [1], would not be expected to be able to be reduced by arsenite oxidase, it nevertheless appears likely that this is the site through which electron transfer would occur.

Table 1. Summary of Data Collection and Refinement Statistics

Crystal Form[a]	$P2_1$	P1[b]
Data Collection		
pH	8.5	6.4
Unit cell dimensions	*a* = 96.7 Å *b* = 114.3 Å *c* = 109.0 Å β = 112.4°	*a* = 90.7 Å *b* = 109.5 Å *c* = 117.6 Å α = 97.7° β = 90.0° γ = 96.4°
Wavelength (Å)	1.746	0.98
Resolution range (Å)	48.2–2.03	17.8–1.64
Number of observations	513,617	930,124
Number of unique reflections	131,366	507,315
Completeness[c] (%)	93.2 (74.9)	93.1 (71.5)
R_{merge}[c] (%)	5.9 (38.9)	5.6 (25.6)
$I/\sigma(I)$[c]	16.4 (2.0)	9.6 (2.9)
Refinement		
Resolution range (Å)	20.0–2.03	17.8–1.64
Number of reflections		
Working set	124,523	481,801
Free set	6,562	25,514
Number of atoms		
Protein	14,524	29,772
Solvent	1,304	4,123
R factor (%)	19.3	15.4
Free R factor (%)	22.3	17.9
Average B factors		
Main chain (Å^2)	26.5	14.6
Side chain (Å^2)	24.8	15.8
Solvent (Å^2)	32.4	27.3
Rmsd from ideality		
Bond distance (Å)	0.02	0.03
Bond angle (°)	1.8	2.6

[a] Statistics are shown only for the two data sets used in the final refinements.
[b] The P1 crystal used in the final refinement was a $HgCl_2$ derivative.
[c] Figures in parentheses refer to the data in the outermost shell.

Biological Implications

Arsenite oxidase from *Alcaligenes faecalis* and related organisms catalyzes an important step in the environmental arsenic cycle, converting the highly toxic arsenite to the relatively more innocuous arsenate. In addition to its principal mode of action in binding to vicinal dithiols such as glutathione, thereby disrupting intracellular redox-homeostasis, arsenite inhibits xanthine oxidase by specifically binding to its molybdenum center. Earlier X-ray absorption spectroscopy (XAS) studies have demonstrated that inhibition involves coordination of arsenite to an active site Mo-SH in the reduced form of the enzyme [19].

From the present structure, it is evident that arsenite oxidase avoids such inhibition by utilizing a fundamentally different type of molybdenum center than is seen in xanthine oxidase and other molybdenum hydroxylases, with the metal coordinated by two equivalents of the pyranopterin cofactor and lacking a Mo-SH in the reduced form. Indeed, with its bis-enedithiolate coordination, the active site of arsenite oxidase is clearly a member of another distinct group of molybdenum-containing enzymes, those from bacterial sources that catalyze oxygen atom transfer reactions. There are three subgroups

to this family of enzymes, possessing, respectively, a serine, cysteine, or selenocysteine ligand from the protein itself. The active site seen here for arsenite oxidase, with a hydroxyl ligand in the position occupied in other members of this family by a protein ligand, represents the first example of a new subgroup of this family of enzymes that lacks any protein ligand to the metal.

The present structure provides a basis for further investigation of the roles of specific amino acid residues and of the catalytic cycle. We are now working toward obtaining additional structures, including the oxidized and inhibitor-bound forms.

Experimental Procedures

Protein Purification

Arsenite oxidase was purified from *Alcaligenes faecalis* NCIB 8687 as described previously [1].

Crystallization and Crystallographic Data

Crystals of arsenite oxidase in two space groups were grown at 4°C using the hanging drop vapor diffusion method. Crystals with space group P1 were obtained by equilibration against a reservoir solution containing 12% polyethylene glycol (PEG) 17,500 and 0.1 M MES (pH 6.4). $HgCl_2$ derivatives were obtained by cocrystallizing with 0.25 mM $HgCl_2$ and equilibrating against a reservoir solution containing 9% PEG 6,000, 10 mM $CoCl_2$, 0.1 M $CaCl_2$, and 0.1 M MES (pH 6.4) or 9% PEG 10,000, 10 mM $CoCl_2$, 0.1 M $CaCl_2$, and 0.1 M MES (pH 6.4). An ethylmercurithiosalycilate (EMTS) derivative was obtained by cocrystallizing with 2 mM EMTS and equilibrating against a reservoir solution containing 10% PEG 17,500, 0.4 M $CaCl_2$, and 0.1 M MES (pH 6.4). Crystals, which are brown colored due to the bound iron-sulfur clusters, grew as flat plates with typical dimensions of 0.3 mm × 0.3 mm × 0.05 mm in approximately 2 weeks. The crystals were invariably epitaxially twinned, consisting of two or more domains layered along the short crystal dimension. Although no perfectly single crystals were ever obtained, data could be collected from selected crystals with very small twinning fractions or from crystals that had been tapered by the surface of the hanging drop, exposing an untwinned domain at the edge of the crystal.

A single crystal with space group $P2_1$ was obtained by equilibration against a reservoir solution containing 15% PEG 6,000 and 0.1 M Tris-HCl (pH 8.5). This crystal had dimensions of 0.3 mm × 0.3 mm × 0.2 mm and grew in approximately 2 weeks. Subsequent attempts to reproduce this form have failed.

As the crystals were sensitive to changes in temperature and rapidly lost diffracting power at room temperature, all crystals were flash cooled in the cold room prior to data collection. This was accomplished by soaking each crystal in a cryoprotectant solution consisting of a 1:3 or 1:4 mixture of ethylene glycol:appropriate reservoir solution for a few seconds before freezing in liquid nitrogen. During data collection, the crystals were maintained at approximately 110 K in a stream of nitrogen gas.

Data Collection

Monochromatic diffraction data were collected from the frozen P1 crystals using MAR345 imaging plate detectors on beam lines 7–1 ($\lambda = 1.08$ Å) and 9–1 ($\lambda = 0.98$ Å, 0.80 Å) at the Stanford Synchrotron Radiation Laboratory (SSRL). Data sets were collected on a native crystal and three Hg derivatives. Anomalous diffraction data around the Fe edge for MAD phasing were collected from the frozen $P2_1$ crystal using an ADSC Quantum4 CCD detector operated by the software package BLU-ICE on SSRL beam line 9–2. The P1 crystals had average unit cell dimensions of $a = 90.7$ Å, $b = 109.5$ Å, $c = 117.6$ Å, $\alpha = 97.7°$, $\beta = 90.0°$, and $\gamma = 96.4°$, and the $P2_1$ crystal had unit cell dimensions of $a = 96.7$ Å, $b = 114.3$ Å, $c = 109.0$ Å, and $\beta = 112.4°$. All data were processed using MOSFLM [20] and the SCALA and TRUNCATE programs from the CCP4 program package [21]. A summary of the data collection and reduction statistics is given in Table 1.

Structure Solution and Refinement

Structure solution by the MIR method was initially attempted using the P1 crystal data. Four heavy-atom sites were located in the EMTS derivative by inspection of the 16 peaks in an anomalous Patterson map. In a similar way, eight heavy-atom sites were located in the $HgCl_2$ derivatives, of which four corresponded to the EMTS positions. Four further minor sites in the $HgCl_2$ derivatives were identified in log-likelihood gradient maps calculated using SHARP [22]. The Mo and Fe atoms of four independent arsenite oxidase molecules were located in a phased anomalous difference map, appearing as four similar sets of a single peak, a pair of peaks and a triangle of peaks, assigned as the Mo atom, a [2Fe-2S] cluster, and a [3Fe-4S] cluster, respectively. Despite this success, however, this approach did not ultimately yield interpretable maps.

The structure was finally solved by the MAD method using the $P2_1$ crystal data. Six anomalous scattering sites were located by analysis of an anomalous Patterson map that was calculated using the peak wavelength data. As it proved difficult to locate the individual metal atoms using the $P2_1$ crystal data alone, the Fe and Mo atom positions were identified using the known coordinates from the P1 structure. Four possible arrangements of the Mo and Fe atoms, combining the two possible hands of the known P1 metal coordinates and the two possible hands of the $P2_1$ metal site coordinates, were generated by superposing the Mo and Fe atoms taken from the P1 model onto the $P2_1$ sites. For each arrangement, phases were calculated using MLPHARE [21, 23] and improved with DM [21, 24], applying solvent flattening, histogram matching, and noncrystallographic symmetry averaging. The correct solution yielded excellent density with right-handed helices. The map was further improved with wARP [25], and the protein backbone was then traced using the automated procedure in the same package. The autotraced model consisted of 1814 residues in 2 independent molecules, or approximately 95% of the total. Finally, the Mo and Fe atoms, inorganic sulfide atoms, and molybdopterin guanine dinucleotide (MGD) cofactors were added by hand to complete the initial $P2_1$ model. An initial P1 model was generated by superposing the metal sites in one $P2_1$ molecule onto the known P1 Mo and Fe coordinates.

The P1 and $P2_1$ models were refined using alternating cycles of optimization with CNS [26] and manual rebuilding in σ_A-weighted $2mF_o$-DF_c and annealed mF_o-DF_c and log-likelihood gradient omit maps using the program XFIT from the XtalView package [27]. In the early stages of the refinement of the $P2_1$ model, combined maps incorporating the experimental phase information were also used. All maps were calculated using CNS. The P1 model was refined using the highest resolution data available, a 1.64 Å data set from one of the $HgCl_2$ derivatives. In the $P2_1$ case, because of concerns about crystal degradation, the 2.03 Å inflection point data set first collected from the crystal was used. For each data set, 5% of the reflections were randomly selected for the calculation of R_{free} and were omitted from all other calculations. For both models, the CNS refinement used a maximum-likelihood target and included an overall anisotropic temperature factor and a flat bulk-solvent correction. This permitted all data to be used. In the final stages of refinement, solvent molecules were added to the models using the automated procedure in CNS. All solvent molecules in the final models correspond to peaks of at least 4σ in log-likelihood gradient omit maps and have real-space correlation coefficients of at least 0.71. Refinement statistics are given in Table 1.

As the amino acid sequence was initially only known for the 18 residues at the N terminus of the large subunit [1], each residue was assigned manually during the course of refinement on the basis of (1) map density for the four independent molecules in the P1 form and the two independent molecules in the $P2_1$ form and (2) hydrogen bonding patterns. Further evidence for the identity of many cysteine, methionine, aspartate, and glutamate residues was obtained from phased (F_o-F_o) maps comparing subsequent $P2_1$ data sets with the first. For approximately half of the methionine and cysteine residues, these maps clearly showed radiation-induced loss of the sulfur atom, and for approximately half of the aspartate and glutamate residues, the maps showed radiation-induced decarboxylation [28]. Subsequently, the models were corrected using a DNA sequence for residues 136–791 of the large subunit. Based on the agreement between the two sequences for this portion of the protein, the 284

residues for which no sequence has been determined, comprising residues 19–135 and 792–825 of the large subunit and all of the Rieske subunit, are expected to incorporate approximately 30 incorrect amino acids, of which half will be Asp/Asn or Glu/Gln substitutions, while the remainder will principally be truncations of long side chains such as Lys and Glu to Ala or Ser.

The final models were analyzed using the program PROCHECK [29]. All parameters are within acceptable limits. For the *P*1 model, the Ramachandran plot shows 89.4% of protein residues in the most favored regions, 9.9% in the additional allowed regions, 0.6% in the generously allowed regions, and only 0.1% in the disallowed regions, comprising the Ala128 residues in each of the four independent long chains. All four of these residues are in excellent density. For the $P2_1$ model, 88.9% of the residues are in the most favored regions, 10.6% are in the additional allowed regions, 0.6% are in the generously allowed regions, and there are no residues in the disallowed regions.

Accession Numbers

The atomic coordinates of the P1 and $P2_1$ forms of arsenite oxidase have been deposited in the Protein Data Bank [30] with accession codes 1G8K and 1G8J, respectively.

Acknowledgments

We would like to thank Aina Cohen, Timothy McPhillips, Michael Soltis, and Eric de La Fortelle for helpful discussions. This work is based upon research conducted at the Stanford Synchrotron Radiation Laboratory (SSRL). SSRL is funded by the Department of Energy (BES, BER) and the National Institutes of Health (NCRR, NIGMS).

Received: August 21, 2000
Revised: November 20, 2000
Accepted: December 1, 2000

References

1. Anderson, G.L., Williams, J., and Hille, R. (1992). The purification and characterization of arsenite oxidase from *Alcaligenes faecalis*, a molybdenum-containing hydroxylase. J. Biol. Chem. *267*, 23674–23682.
2. Hille, R. (1996). The mononuclear molybdenum enzymes. Chem. Rev. *96*, 2757–2816.
3. Kisker, C., Schindelin, H., and Rees, D.C. (1997). Molybdenum-cofactor-containing enzymes: structure and mechanism. Annu. Rev. Biochem. *66*, 233–267.
4. Bard, A.J., Parson, R., and Jordon, J. (1985). Standard Potentials in Aqueous Solutions (New York: Marcel Dekker).
5. Boyington, J.C., Gladyshev, V.N., Khangulov, S.V., Stadtman, T.C., and Sun, P.D. (1997). Crystal structure of formate dehydrogenase H: catalysis involving Mo, molybdopterin, selenocysteine, and an Fe_4S_4 cluster. Science *275*, 1305–1308.
6. Dias, J.M., et al., and Romao, M.R. (1999). Crystal structure of the first dissimilatory nitrate reductase at 1.9 Å solved by MAD methods. Structure *7*, 65–79.
7. Schindelin, H., Kisker, C., Hilton, J., Rajagopalan, K.V., and Rees, D.C. (1996). Crystal structure of DMSO reductase: redox-linked changes in molybdopterin coordination. Science *272*, 1615–1621.
8. Schneider, F., Löwe, J., Huber, R., Schindelin, H., Kisker, C., and Knäblein, J. (1996). Crystal structure of dimethyl sulfoxide reductase from *Rhodobacter capsulatus* at 1.88 Å resolution. J. Mol. Biol. *263*, 53–69.
9. Czjzek, M., Dos Santos, J.P., Pommier, J., Giordano, G., Méjean, V., and Haser, R. (1998). Crystal structure of oxidized trimethylamine *N*-oxide reductase from *Shewanella massilia* at 2.5 Å resolution. J. Mol. Biol. *284*, 435–447.
10. Lesk, A.M. (1995). NAD-binding domains of dehydrogenases. Curr. Opin. Struct. Biol. *5*, 775–783.
11. Shindyalov, I.N., and Bourne, P.E. (1998). Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng. *11*, 739–747.
12. Lu, G. (2000). TOP: A new method for protein structure comparisons and similarity searches. J. Appl. Crystallogr. *33*, 176–183.
13. Iwata, S., et al., and Jap, B.K. (1998). Complete structure of the 11-subunit bovine mitochondrial cytochrome bc_1 complex. Science *281*, 64–71.
14. Iwata, S., Saynovits, M., Link, T.A., and Michel, H. (1996). Structure of the water soluble fragment of the 'Rieske' iron-sulfur protein of the bovine heart mitochondrial cytochrome bc_1 complex determined by MAD phasing at 1.5 Å resolution. Structure *4*, 567–579.
15. Carrell, C.J., Zhang, H., Cramer, W.A., and Smith, J.L. (1997). Biological identity and diversity in photosynthesis and respiration: structure of the lumen-side domain of the chloroplast Rieske protein. Structure *5*, 1613–1625.
16. Kauppi, B., et al., and Ramaswamy, S. (1998). Structure of an aromatic-ring-hydroxylating dioxygenase – naphthalene 1,2-dioxygenase. Structure *6*, 571–586.
17. Volbeda, A., Charon, M.H., Piras, C., Hatchikian, E.C., Frey, M., and Fontecilla-Camps, J.C. (1995). Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. Nature *373*, 580–587.
18. Page, C.C., Moser, C.C., Chen, X., and Dutton, P.L. (1999). Natural engineering principles of electron tunneling in biological oxidation-reduction. Nature *402*, 47–52.
19. Cramer, S.P., and Hille, R. (1985). Arsenite-inhibited xanthine oxidase - determination of the molybdenum-sulfur-arsenic geometry by EXAFS. J. Am. Chem. Soc. *107*, 8164–8169.
20. Leslie, A.G.W. (1999). Integration of macromolecular diffraction data. Acta Crystallogr. D *55*, 1696–1702.
21. Collaborative Computational Project No 4. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760–763.
22. de La Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. In Methods Enzymology, Macromolecular Crystallography, R.M. Sweet, and C.W. Carter, eds. (New York: Academic Press), pp. 472–494.
23. Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. In Isomorphous Replacement and Anomalous Scattering Proceedings of the CCP4 Study Weekend, 25–26 January 1991, W. Wolf, P.R. Evans, and A.G.W. Leslie, eds. (Daresbury, UK: Science and Engineering Research Council), pp. 80–86.
24. Cowtan, K. (1999). Error estimation and bias correction in phase-improvement calculations. Acta Crystallogr. D *55*, 1555–1567.
25. van Asselt, E.J., Perrakis, A., Kalk, K.H., and Lamzin, V.S. (1998). Accelerated X-ray structure elucidation of a 36 kDa muramidase/transglycosylase using wARP. Acta Crystallogr. D *54*, 58–71.
26. Brünger, A.T., et al., and Warren, G.L. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D *54*, 905–921.
27. McRee, D.E. (1999). XtalView/Xfit - A versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. *125*, 156–165.
28. Burmeister, W.P. (2000). Structural changes in a cryo-cooled protein crystal owing to radiation damage. Acta Crystallogr. D *56*, 328–341.
29. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
30. Berman, H.M., et al., and Bourne, P.E. (2000). The protein data bank. Nucleic Acids Res. *28*, 235–242.
31. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structure. J. Appl. Crystallogr. *24*, 946–950.
32. McDonald, I.K., and Thornton, J.M. (1994). Satisfying hydrogen bonding potential in proteins. J. Mol. Biol. *238*, 777–793.